Master's Thesis

Multi-Scale Approaches for Pathology Image Analysis

WonJae Hong

Department of Electrical and Computer Engineering
(Electrical Engineering)

Graduate School of UNIST

2020

Multi-Scale Approaches for Pathology Image Analysis

WonJae Hong

Department of Electrical and Computer Engineering
(Electrical Engineering)

Graduate School of UNIST

Multi-Scale Approaches for Pathology Image Analysis

A thesis/dissertation
submitted to the Graduate School of UNIST
in partial fulfillment of the
requirements for the degree of
Master of Science

WonJae Hong

12 / 04 / 2019
Approved by

Advisor
Se Young Chun

Multi-Scale Approaches for Pathology Image Analysis

WonJae Hong

This certifies that the thesis/dissertation of WonJae Hong is approved.

4 December 2019

Advisor: Se Young Chun

Won Ki Jeong: Thesis Committee Member #1

Jeong Hwan Jeon: Thesis Committee Member #2

Abstract

The multi-scale approach has been applied to deep learning methods for image processing such as image segmentationand image super-resolution. So we researched the impact of using the multi-scale information in terms of image segmentation and image generation in this thesis. In lesion segmentation, Segmentation of skin lesions is often done using image segmentation as an important pre-processing step before skin melanoma images are analyzed. With the advent of U-Net in the field of medical image segmentation, it is possible to segment skin lesions of superior quality. We would like to propose a structure that can use multi-scale information more effectively based on U-Net. Specifically, we devised a deep neural network that combines various U-Nets operating in multi-scale to obtain final segmentation results through U-Nets that perform final skin lesion segmentation. We have examined the advantages and disadvantages of the proposed multi-scale method by combining experiments at various resolutions.

In image genetation research, We made various attempts to generate pathological images using SinGAN with a multiscale pyramid pipeline structure. During training, SinGAN are learning to generate down-sampled input image in the initial steps, so it is focused on global features. And, as scale goes up, it is focused on fine areas. We applied them to various image classifications using GAN-generated images. Through patient classification experiments, we attempted to identify the effect of preventing the leakage of personal information of concern from pathological images and identified about 30% degradation in classification performance. And through the Tumor classification experiment, we checked the possibility of using the GAN-generated image as data to train the model, and confirmed its value as training data by having the same performance as real pathology images. However, since the two experiments are conducted on a small sample, additional experiments are required after collecting more data.

Contents

List of Figures

List of Tables

CHAPTER I

Introduction

1.1 Aim of this research

The aim of this research is to investigate a way of performance enhancement of deep learning models for image processing using multi-scale images. The multi-scale approach has been applied to deep learning methods for image processing such as image segmentation [12, 13] and image super resolution [14].If image is converted in various ways, even if it's the same image, Deep Learning Network can recognize each converted image as a different object. In particular, since images of multi-scale can give different information, image processing methods performed at multi-scale method use different features that can be obtained at various scales. Therefore, if we convert the images into various scales and train them using the same network, the predicted results from the network model will be different even if you use same network and same image. So we researched the impact of using the multi-scale information in terms of image segmentation and image generation in this thesis.

1.2 Multi-scale approach for Skin lesion segmentation.

Skin cancer is one of the major health problems in the United States with 5 million diagnoses each year. If diagnosed early, however, survival exceeds 95%. That is why early diagnosis of skin cancer is so important. There are a variety of image analysis methods available to enable

automatic diagnosis of melanoma from skin images. Among the various image analysis methods, skin lesion image segmentation is an important task to improve overall analysis performance. We propose a network that uses the multi-scale method for image analysis to accurately segment lesion areas in skin lesion images. For this, We propose a network that extends the most widely used U-Net for medical image analysis and segmentation to combine and use the results of training images of different sizes.

1.3 Multi-scale approach for Pathology Image Generation.

Recently, research on medical image analysis has been actively conducted using deep learning. Various analysis methods have been developed, such as detecting cancer metastasis [15,16], segmentation of tumer and classifying the size of the tumer [17, 18], and classifying and diagnosing disease through pathological images [19, 20]. The main reason for this research is that it is hard to find abnormal 100x100 pixels tissue in the 100,000 x 100,000 pixel WSI image. Doctors spend a lot of time analyzing images, and sometimes several doctors spend a lot of time analyzing images. Medical image analysis through deep learning will help doctors make efficient and accurate diagnosis.

However, one of the difficulties of researching deep learning using medical imaging is that data is difficult to obtain. Datasets for general image processing are provided by many institutions, and we can easily obtain them. However, in a hospital or a medical institution that collects medical image data, a review for protecting patient privacy is essential to distribute the collected medical data. If a researcher has difficulty obtaining data because of this procedure, the research process will be delayed.

So in this research, we tried to solve the problem of data acquisition using fake image generation, similar to the actual medical image using Generative Adversarial Networks(GAN) [1].

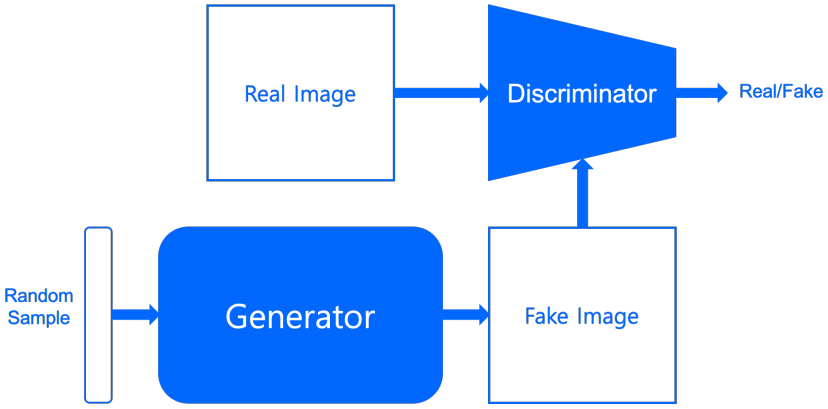

Figure 1.1: Generative Adversarial Network [1]

GAN is composed of two models, one is the Generative model and the other is the Discriminative model. Generative models generate fake data, and Discriminative model classify fake and real using fake images and real images generated from the Generative model. So, we tried to generate fake images similar to the actual medical images by using SinGAN [10] using a multi-scale approach among various GANs, and to verify whether the fake images function like the real images.

CHAPTER II

Multi-scale Approach for Skin Lesion Segmentation

2.1 Background and Related Works

2.1.1 U-Net

U-Net [2] is a convolutional neural network developed for medical image segmentation and has shown to be effective in research on image segmentation. The network is based on a fully convolutional network [21], and the architecture has been modified and extended to work with less training images and generate more accurate segments.

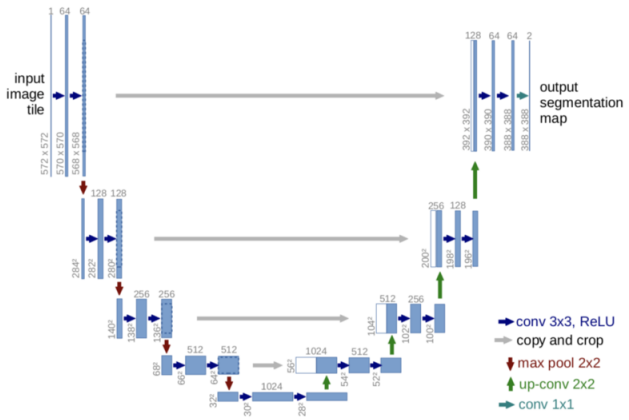

Figure 2.1: Architecture of U-Net [2] for medical image segmentation.

The architecture of U-net network consists of two parts. One is Contracting Path(left) corresponding to Convolution Encoder, and the other is Expanding Path(right) corresponding to Convolution Decoder. Contracting path is mainly composed of 3x3 convolutional layers. Two 3x3 convolutional layers and max pooling are used for each convolutional block. The feature that has undergone max pooling doubles the number of channels compared to the feature of the previous block. Expanding path is also composed of 3x3 convolutional layers. At end of each block, 2x2 up-convolutional layer is used that makes the number of channels in half. On top of block in expanding path, The first feature of each block is cropped before Max-Pooling in the Contracting Path and combined with the up-convoution feature from the previous block. Final layer uses 1X1 convolutional layer to classify each pixel into two classes and generates a segmentation mask.

In our research, U-Net is used as a baseline. Using U-Net, we generate segmentation masks of lesion images at multi-scale. Then, we propose a method of properly using the mask obtained through the image of each scale to improve the performance.

2.1.2 Data distillation structure for ensemble of multi-scale mask

When we design deep learning models, ensemble is is easy and effective ways to improve performance. So, ensemble is used in many competitions. Ensemble is a method of combining the results from different initializations in the same model or from different neural networks.

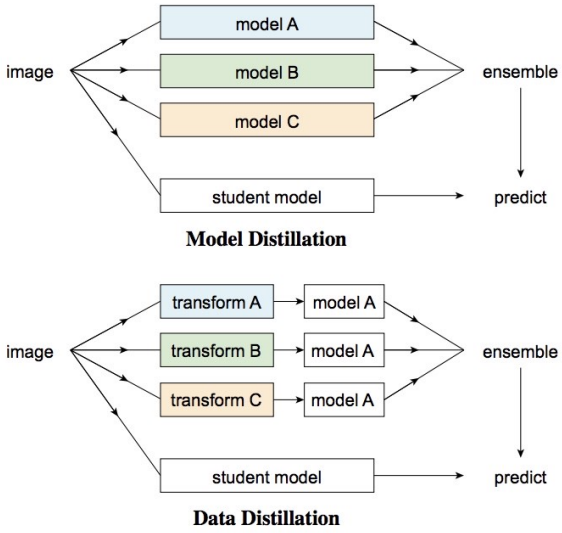

Figure 2.2: Architecture of Model distillation [3] and Data distillation [4]. Model distillation generates ensemble from different model ans same input. Data distillation generates ensemble from different input and same model.

Figure 2.2 shows architecture of Model distillation [3] and Data distillation [4]. In our research, we adopt architecture of Data distillation. Originally, distillation method is designed for transfer ensemble from bigger neural networks to smaller neural network that is called student network. But we use upper architecture of data distillation. At upper architecture of data distillation, input image is transformed and train the model through the transformed image. After that, ensemble is generated by combining the outputs from the trained model of each transformed image. We transformed skin lesion images on various scales and trained each image on U-Net. We designed a method to generate an ensemble using the output of U-Net, which trained images of various scales.

2.2 Method

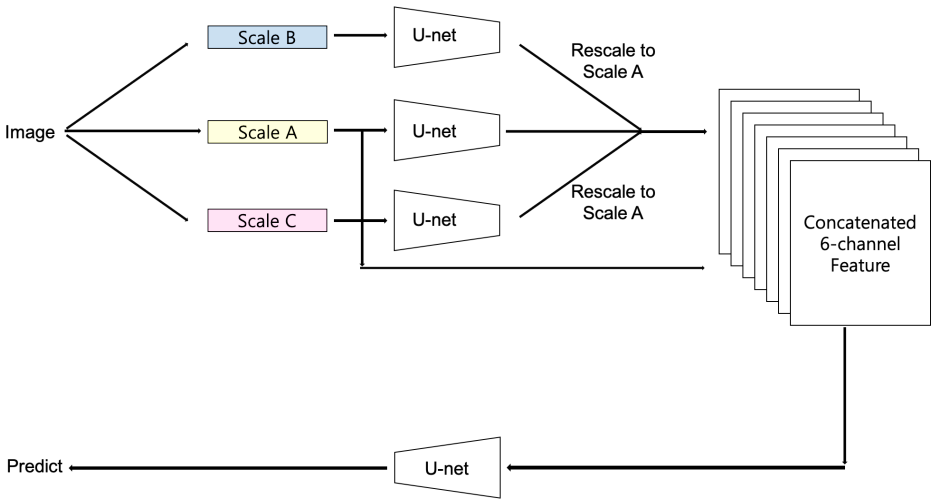

Figure 2.3: Architecture of image segmentation model using multi-scale skin lesion images.

We designed a multi-scale based skin lesion image segmentation network using U-Net. Figure 2.3 is architecture of our proposed network. The network consists of a combination of U-Net and the upper part of the data distillation structure. The input skin lesion images are converted into 3 different scales each 256x192, 192x128, 128x96 pixels. Three independent U-Nets are trained using three transformed (rescaled) images. For we generate ensemble, we rescale the trained independent 3 U-Net 1-channel output segmentation masks into one of three scales (scale can be selected) and three rescaled segmentation mask are concatenated into three channel features. After that, we concatenate the 3-channel features with the input image to genetate a feature with a total of 6-channels. Finally, the 6-channel feature is again used as training data for the U-Net, and the result of 1-channel output segmentation becomes the final output, that is, the image segmentation result.

2.3 Experiments

2.3.1 Dataset

In this experiment, Training Dataset and Test Dataset used the training data provided by ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection Task1: Lesion segmentation hosted by MICCAI'18 [5, 6].

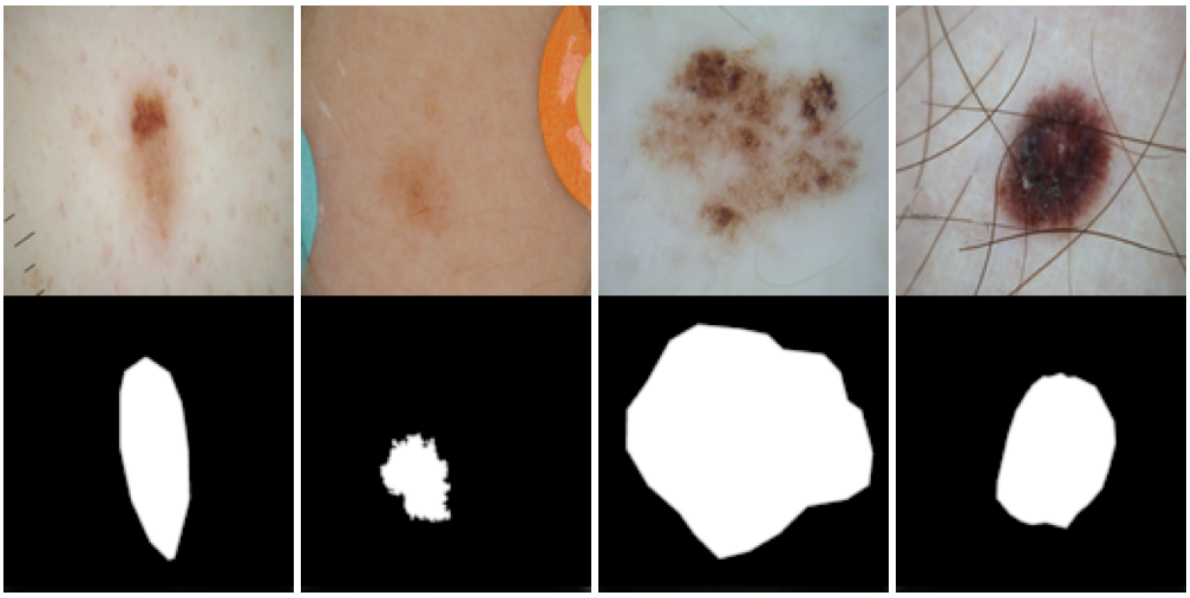

Figure 2.4: Dataset provided by ISIC 2018:Skin Lesion Analysis Towards Melanoma Detection. [5, 6]

The total number of images consists of 2594 skin lesion images and their corresponding ground truth. Of the 2594 images, 2200 images were used as a trainset for training the proposed model. Trainset is augmented to 8800 images by using flipping in three ways: top, bottom, left, right, top and bottom. The 394 images were used as a testset to verify the performance of the proposed model. For model training using multi-scale images, images were converted to three different scales: 256x192, 192x128, and 128x96. The original image always has higher resolution than these three scales.

2.3.2 Experiment Settings

In the training phase of the model using the proposed method, the batch size was set to 16. We also used Adam optimizer [22] as the optimization function and set the learning rate of the optimization function to 0.0001.We use binary cross-entropy as the loss function. The implementations of them are performed on Keras [23].

$$CE(x) = -\frac{1}{N}\sum_{i=1}^{N} y_i \log\left(h(x_i;\theta)\right) + (1-y_i)\log\left(1-h(x_i;\theta)\right) \tag{II.1}$$

2.3.3 Evaluation Metric

In this experiment we used the Jaccard Index (or IOU) as evaluation metric. We calculated the Jaccard Index scores for the predicted and ground truth masks for each skin lesion image, and calculated their average.

$$J(X,Y) = \frac{|X \cap Y|}{|X \cup Y|} = \frac{|X \cap Y|}{|X| + |Y| - |X \cap Y|} \quad , \quad 0 \leq J(X,Y) \leq 1 \tag{II.2}$$

In addition, the scoring method used in ISIC2018 Challenge Task 1 is applied. In ISIC2018, if the jaccard index score is less than 0.65, 0 points are assigned to each image result. By using this approach to exclude results that are meaningless, we intend to give meaning only to results that are practically useful.

2.4 Result

Table 2.1: Average of Jaccard Index scores of existing method and proposed model according to image scale

Method	Image Scale		
	128x96	192x128	256x128
U-NET [2]	0.7748	**0.7809**	0.7787
Proposed Model	0.7710	**0.7818**	0.7757

Table 2.1 compares the Jaccard Index scores of the 128x96, 192x128, and 256x192 images with U-Net only, and the Jaccard Index scores of the proposed model, which outputs the final results on different scales after using all scale information. Using all the scale information in the proposed model and concatenating it on the 192x128 scale shows the best results in all the methods compared. However, concatenating the proposed model with too small (128x96) or too large (256x192) results in worse performance than the basic U-Net performance. Therefore, we find that choosing the appropriate scale when applying the proposed model can help improve the performance of the proposed model.

Table 2.2: Average of Jaccard Index scores that applied ISIC2018 Challenge scoring rule of existing method and proposed model according to image scale

Method	Image Scale		
	128x96	192x128	256x128
U-NET [2]	0.6786	0.6927	**0.7059**
Proposed Method	0.6796	0.7010	**0.7022**

Table 2.2 shows the results of measuring the performance of the proposed model using the method of assigning the Jaccard Index score used in the Lesion Segmentation Challenge of ISIC2018. The above scoring results show that it is particularly advantageous to use relatively high resolution images. For the U-Net results, the highest resolution was the best performance, and the proposed model using the high resolution was able to get better results than the existing U-Net. However, at the highest resolution, the proposed model showed lower performance than the existing U-Net. This can be expected because the proposed model uses low resolution information. This suggests that the proposed model tends to degrade performance in terms of Jaccard index, but increases the accuracy of the prediction results for meaningful predictions.

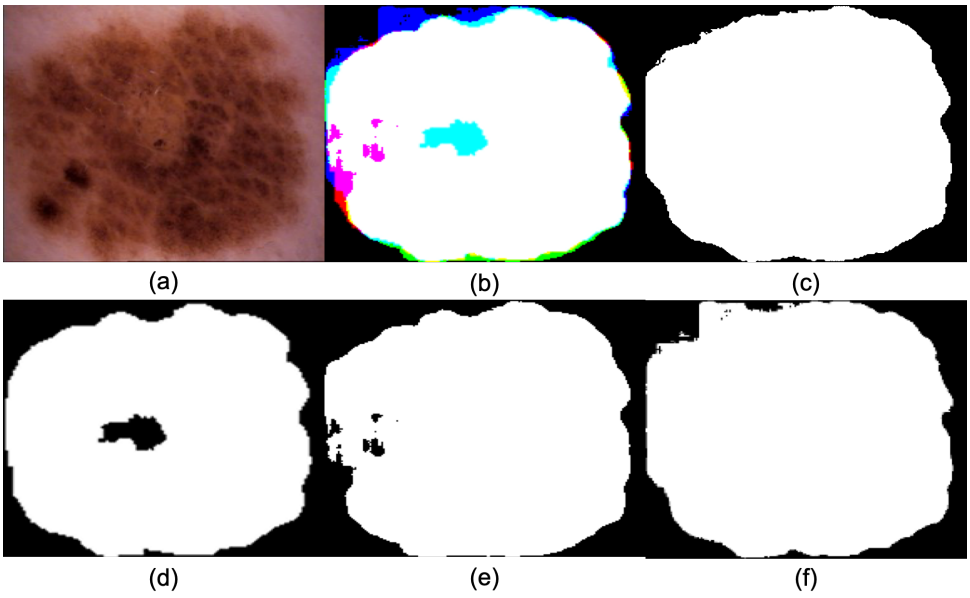

Figure 2.5: (a): Input Image, (b): concatenated feature (c): final prediction, (d): predicted 128x96, (e): predicted 192x128, (f): predicted 256x192

Figure 2.6 shows an input image, concatenated feature map, final segmentation and three intermediate segmentation predictions from three different resolutions. Prediction using a 128x96 scale image shows that the boundary is predicted relatively smoothly, but the center of the lesion area is empty. Predictions through the 192x128 and 256x128 scales show that there is no empty space in the lesion area, but the boundary is not smooth. Using the proposed method,

the final result of concatenating images of different scales shows that the lesion area is predicted almost accurately by taking advantage of both the small resolution and the high resolution.

2.5 Application

Using the proposed model above, we participated in the ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection of MICCAI'18 Task1: Lesion segmentation, earning a Jaccard Index score of 0.726 from 1000 images for performance measurement provided by the organizers, ranking in the middle ranks (61th rank of 112 participants, including duplicate submissions).

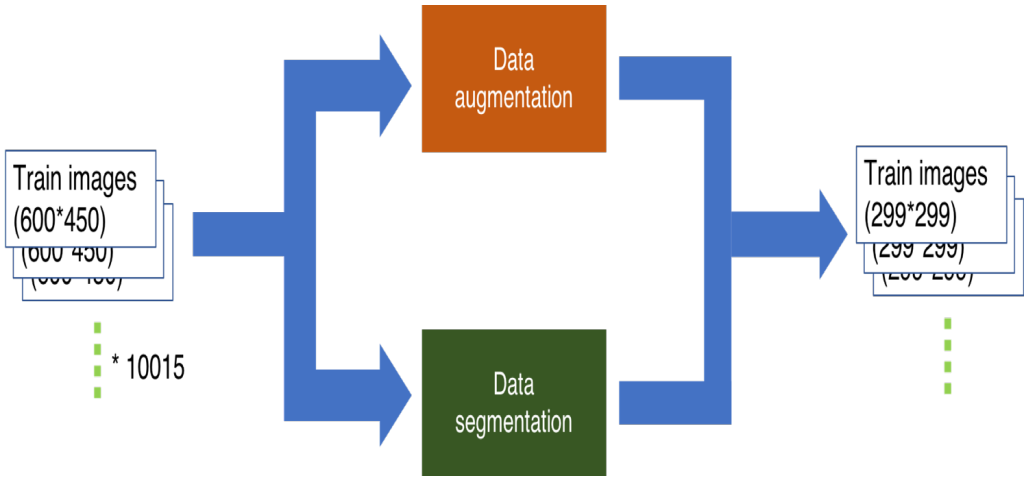

Figure 2.6: Process of concatenating segmentation results and lesion diagnosis training dataset.

Also, we participated in the ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection of MICCAI'18 Task3: Lesion Diagnosis. Using proposed segmentation method, we generate segmentation mask of dataset for lesion classification task. And then, mask is concatenated with original dataset. Using concatenated dataset, we train Inception-V4 [24]. We ranked 73 of 140 participants in Task 3: Lesion Diagnosis (Score is 0.687).

CHAPTER III

Multi-scale Approach for Pathology Image Generation

3.1 Background and Related Works

3.1.1 Generative Adversarial Networks

Since the method of image generation using GAN [1] was first announced, research on GAN has progressed. The limitation of the first published GAN [1] is that learning is unstable to apply in various directions. The model that compensates for the drawbacks is Deep Convolutional GAN(DCGAN) [7].

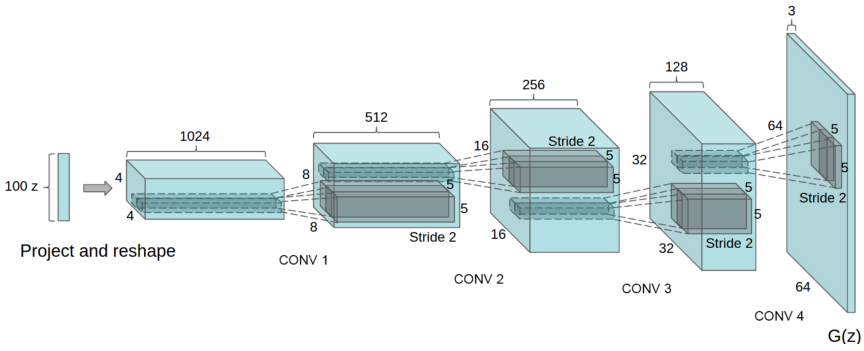

Figure 3.1: Generator Architecture of DCGAN [7]

DCGAN is consist of CNN structures. But, unlike the existing GAN, it does not use a pooling layer. The pooling layer is a layer that reduces the number of unnecessary parameters and selects only important features in many deep learning models, but loses location information of an image. Therefore, DCGAN excluded the pooling layer to preserve location information that is important for image generation. Instead of pooling layer, DCGAN adopts strided convolutions in the Discriminator and fractional-strided convolutions in the Generator. Also, for more stable learning, batch normalization [25] was used. In this way, DCGAN has provided guidelines for stable learning.

Figure 3.2: Left: CycleGAN [8] for image translation, Right: SRGAN [9] for image super-resolution

After DCGAN, there are studies that attempt to learn GAN stably by modifying loss functions such as WGAN [26] and LSGAN [27]. In addition to image generation, GAN was applied to various tasks such as CycleGAN [8] for image translation, SRGAN [9] for image super-resolution In addition to the GAN using images, SEGAN [28] has been proposed to reduce noise in voice recording.

3.1.2 SinGAN

SinGAN [10] is a multi-scale image generative GAN model that train the model using a single image. Previously, GAN has used a method of providing conditional information [29] to input signals or limiting tasks(e.g. super-resolution, image translation) to improve performance. Previously, GAN has used conditional information on input signals or limited tasks to improve performance. In addition, model training requires a large number of data. Recently released InGAN [30] has conducted research to generate multiple images from a single natural image, but uses a conditional method for the input image. Unconditional Single Image GAN [31] studies show good performance in Texture Generation, but not good for natural images.

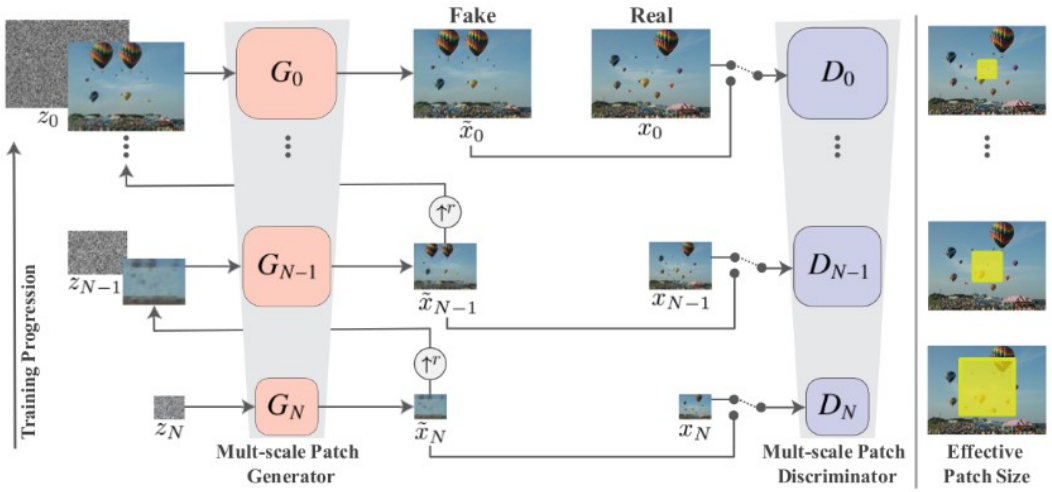

Figure 3.3: SinGAN [10]'s multi-scale pipeline

In the Figure 3.3, x_0 is the training image, which is down-sampled by r times ($r > 1$) as one step goes down. In each step, the generator generates an image with noise and the resulting image generated in the previous step, and the discriminator in that step is trained to distinguish the down-sampled input image from the generated image. In the initial steps, SinGAN are learning to generate down-sampled input image, so it is focused on global features. And, as scale goes up, it is focused on fine areas.

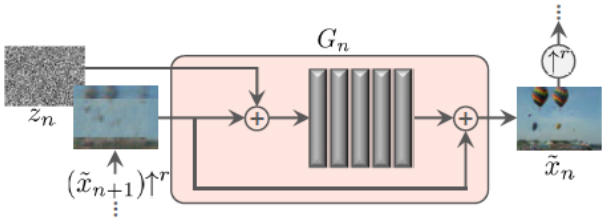

Figure 3.4: n_{th} scale generator of SinGAN [10]

FIgure 3.4 is illustrated generator for each scale in SinGAN. Noise $z_n$is added to the up-sampled image that generated in the previous step. and fed into the fully convolutional network. The feature generated in the fully convolutional network is added to the up-sampled image that generated in the previous step, which is the image generated on the n_{th} scale.

$$\widetilde{x}_n = (\widetilde{x}_{n+1})\uparrow^r + \psi_n(z_n + (\widetilde{x}_{n+1})\uparrow^r) \tag{III.1}$$

Denote by ψ_n is fully convolutional network of the generator has a simple structure that repeats 3x3 Conv-BatchNorm-LeakyReLU 5 times.

SinGAN's Loss function consists of an adversarial term and a reconstruction term.

$$\min_{G_n}\max_{D_n} \quad L_{adv}(G_n, D_n) \quad + \quad \alpha L_{rec}(G_n). \tag{III.2}$$

As adversarial loss, WGAN-GP loss [32] is used, and recontruction loss is

$$L_{rec} \quad = \quad \|G_n(z^*) - x_n\|^2, \quad for \quad n = N \tag{III.3}$$

$$L_{rec} \quad = \quad \left\|G_n(0, \widetilde{x}_{n+1})\uparrow^r)) - x_n\right\|^2, \qquad for \quad n < N \tag{III.4}$$

,where z^* is some fixed noise maps. Reconstruction loss uses squred loss to learn in order to reduce the difference between the pixel generated by the generator and the down-sampled input image at that stage. For learning, the noise setting of each stage injects fixed noise only in the first stage(N stage) and noise is not injected in the remaining stages.

3.2 Method

In this research, we generated a pathological image using SinGAN and conducted two tasks. In the first task, an experiment was conducted to find out that fake pathological images generated by SinGAN are recognized differently from real pathological images. In the second task, we compared the training results of the model with the fake images generated with the GAN with the training results of the model with the real images.

3.2.1 Patient classification

In this task, we conducted an experiment comparing the results of patient classification between real and fake images using ResNet50 overfitted with real images. Patient classification results using ResNet50 [33] overfitted on real images will predict almost all classes. By comparing the results of patient classification using fake images with the results of classification using real images, we tested whether the personal information could be protected through fake images through the difference of classification results.

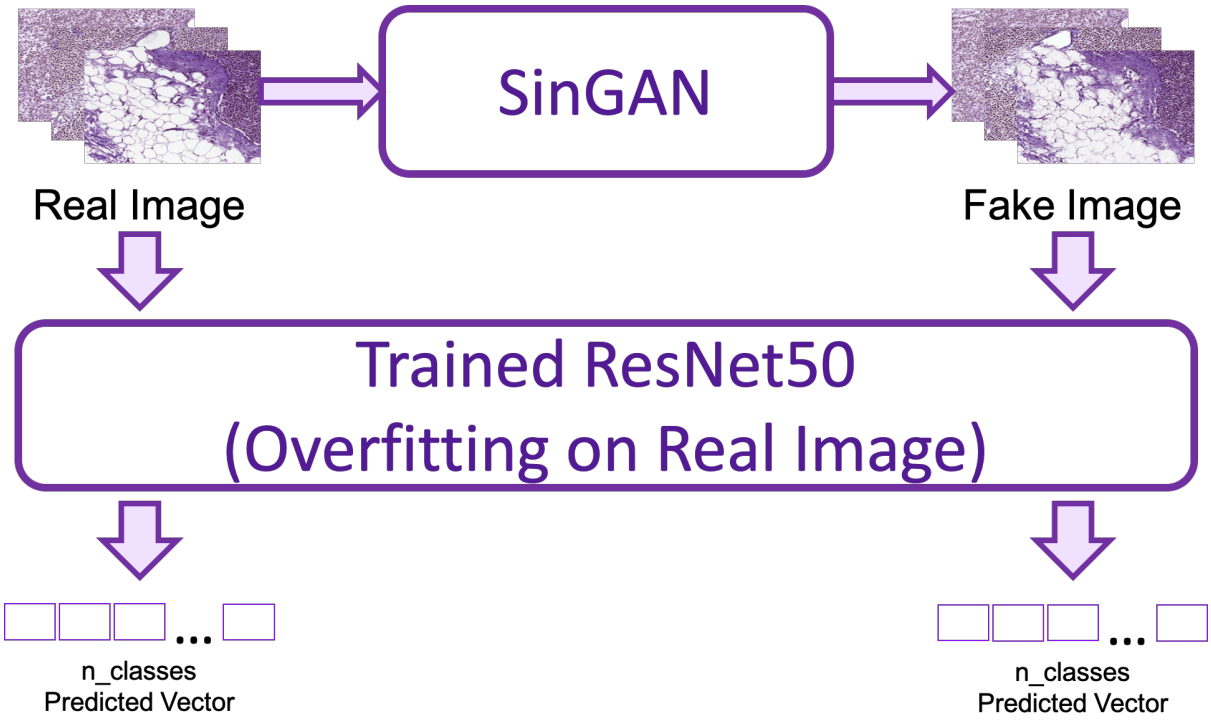

Figure 3.5: Overview of patient classification with real image and fake image

3.2.2 Tumer classification

In the previous task, we focused on how well fake images deceive us. Conversely, this task focuses on how fake images can be used to train deep learning models as much as real images. We divided the real image into trainset and testset. And then, we generated fake images corresponding to real images through SinGAN. The real pathological image training data and the fake pathological image training datawere each trained to classify whether or not tumors were included in the pathological image through ResNet50 [33]. Last, we compared the results of the two distinctions using a test set for each distinct model trained with real and fake images.

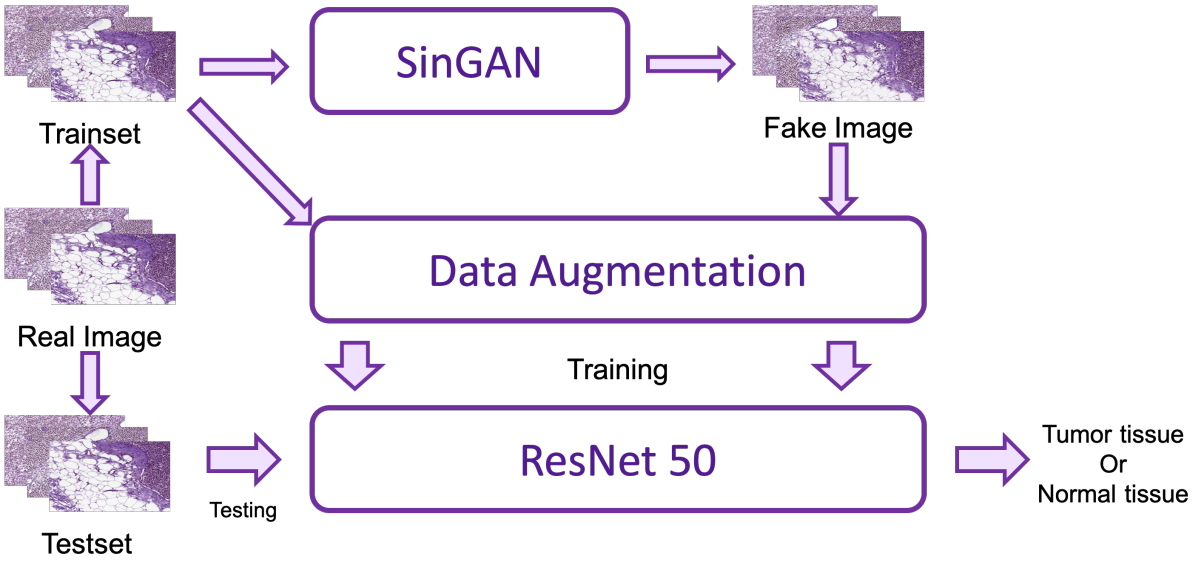

Figure 3.6: Overview of Tumer classification with real image and fake image

3.3 Experiments

3.3.1 Dataset

In this experiment, we used the Whole Slide Image(WSI) provided by the Camelyon17 Challenge [11].

Camelyon17 training dataset has WSI for 99 patient. We selected 20 tumor-negative cases and 20 tumor-positive cases. In each whole slide file(.tif), we chose a Level 3 WSI and divide it into 256x256 patches. And we selected 10 patches per person and used a total of 400 patches as a dataset for SinGAN [10] learning. In our first experiment, we used 400 real patches for 40 people as trainsets and 400 fake patches generated in GAN as testsets. In the second experiment, we set a testset by separating 40 from the 400 real patches. We then constructed the train set

by increasing the remaining 360 real patches and the corresponding 360 fake patches to 11,520 patches each through data augmentation.

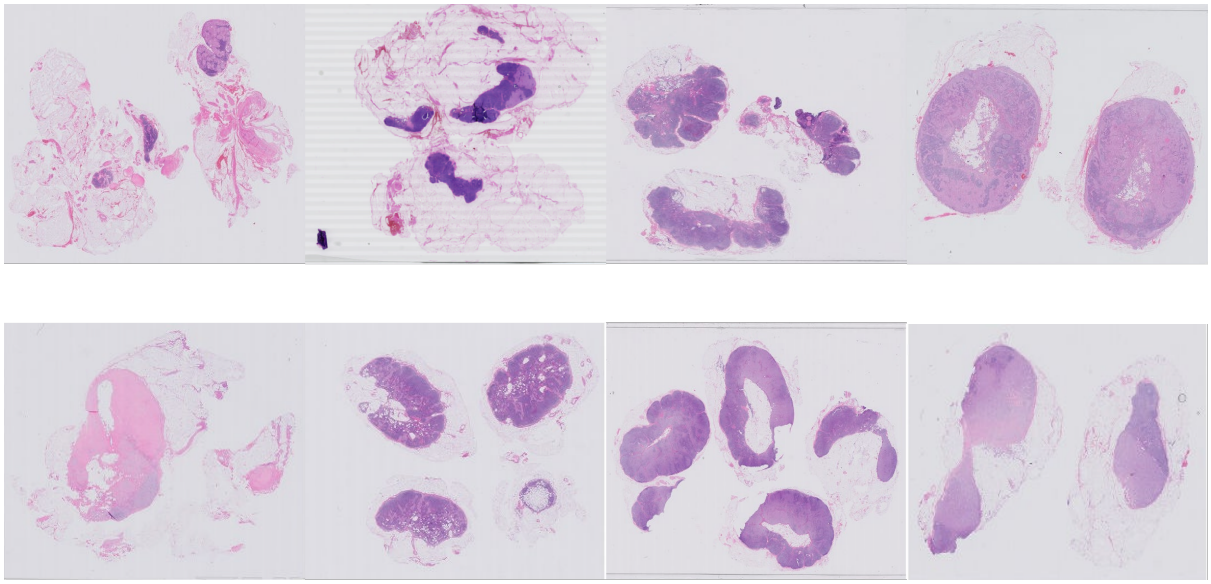

Figure 3.7: Camelyon17 Dataset [11]

3.3.2 Experiment Settings

In the image generation experiment using SinGAN, Adam optimizer [22]is used as the optimizer of generator and discriminator and the learning rate is set to 0.0005 respectively. The number N of scales is set to 5, and the scale factor is set to 0.5. And the hyper parameter (α) of loss function (eq III.2) is set to 15. The implementations of SinGAN are performed on PyTorch [34].

In classification training using ResNet, we used the classifier and Adam optimizer as well, and set the learning rate to 0.0001. The batch size is set to 32 and cross-entropy loss(eq II.1) is used as the loss function. The implementations of SinGAN are performed on Keras [23]. Evaluation metric of classification performance was evaluated by accuracy.

$$Accuracy = \frac{1}{n}\sum_{i=1}^{N} 1(y_{i,true} = y_{i,prediction}) \tag{III.5}$$

3.4 Results

3.4.1 Pathology Image Genaration

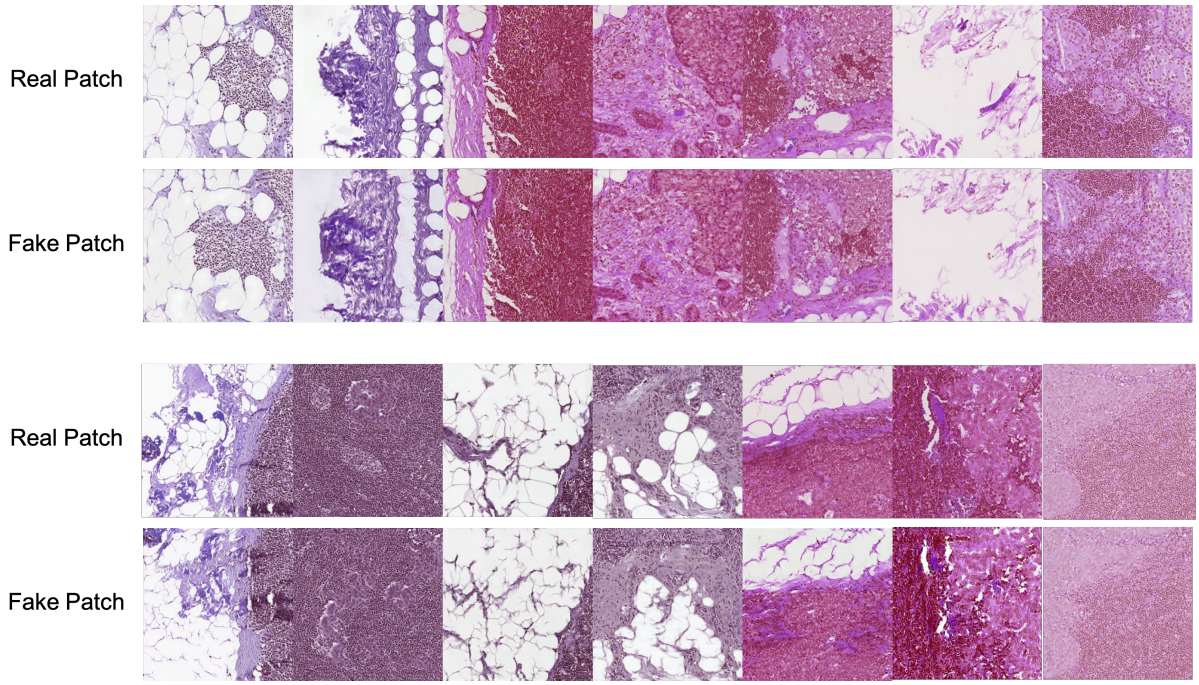

Figure 3.8: Input real image and Generated fake image from SinGAN [10]

Figure 3.8 illustrates a real input image and a fake image generated from SinGAN. Globally, you can see that fake images are generated similar to real input images. This shows that SinGAN's model structure with multi-scale pyramid structure has advantages in preserving overall features. For details, the overall feature is the same, but the pattern of tissues is slightly different, especially for the mesh-shaped or round-shape tissues.

3.4.2 Patient classification

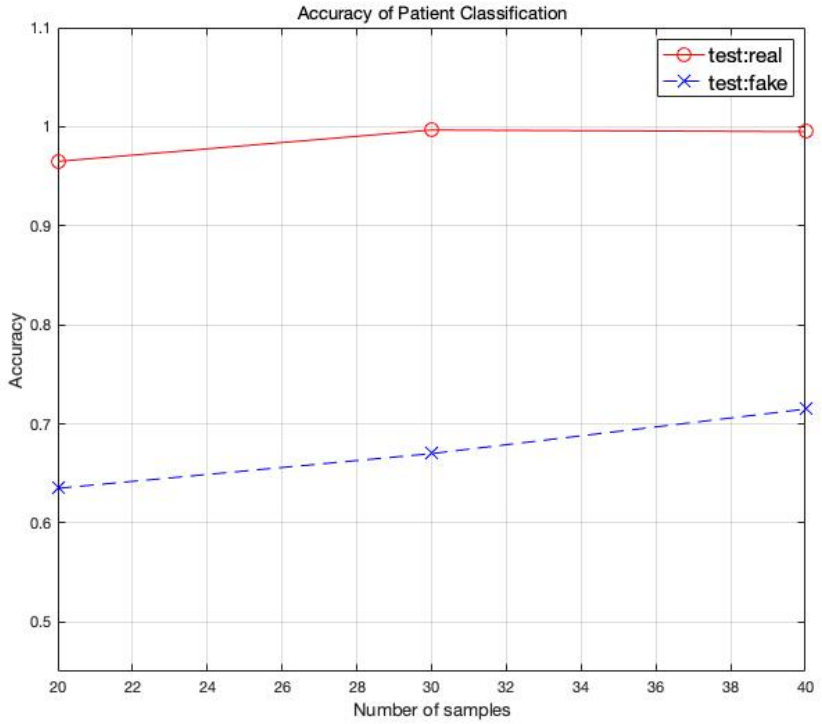

Figure 3.9: Patient Classification Result with Real Image overfited ResNet50

Figure 3.9 is an illustration of the classification accuracy of real and fake images using ResNet50 overfitted with real images. The number of patients tested (number of samples tested) is 20,30,40. As a result of testing the classification performance with real data, the classification accuracy was 96.5% for 20 people, 99.67% for 30 people, and 99.5% for 40 people. The result of testing classification performance with fake images generated through GAN showed 63.5% for 20 people, 67% for 30 people, and 71.5% for 40 people. The difference between the accuracy tested with real images and the accuracy tested with fake images was 33%, 32.67%, and 28%, respectively. The drop in accuracy is an average of 31.2%.

3.4.3 Tumor classification

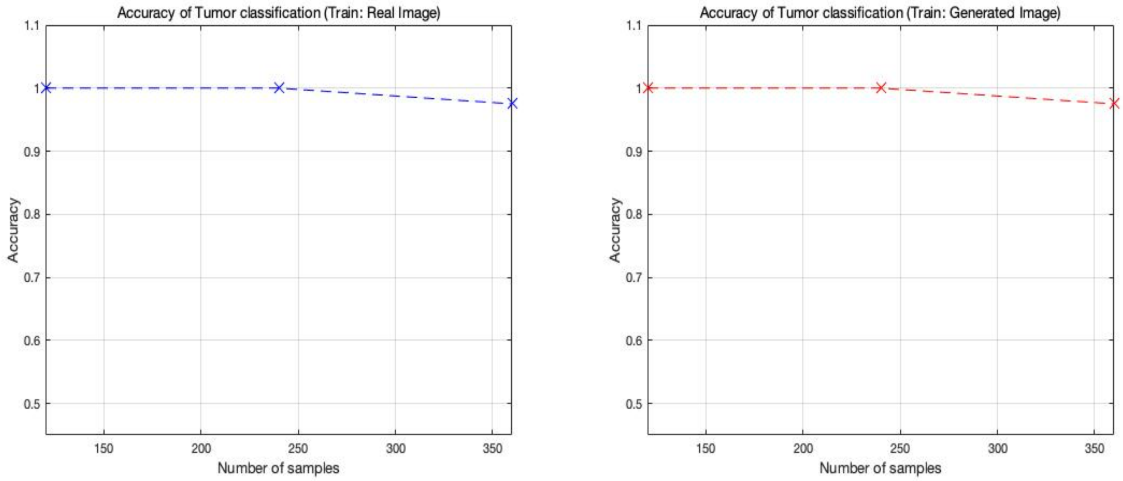

Figure 3.10: Tumer Classification result with real pathology patch(Left) and generated pathology image through GAN(Right)

In this experiment, we conducted an experiment on whether the model training was performed using fake pathology images generated through GAN, and the same effect was obtained when training the model through real pathology images. We made trainsets the actual pathological images of 360 images and the corresponding fake images 360 images, and sequentially trained 120, 240, and 360 tumors individually. Using real pathology images as training data, we tested 120 and 240 models on 40 test sets, and all 40 images accurately classified tumor-containing images and when used 360 images, we obtained a classification accuracy of 97.5%. In addition, when using the image generated through GAN as training data, the model showed the same performance as the training results using the real image.

3.5 Discussion

In our patient classification experiments, we identified a 31.2% reduction in classification performance when classifying fake images generated in GAN using ResNet overfitted with real images. In this result, we identified the possibility that even if an image created with a GAN is leaked, it may not be able to identify the corresponding patient's personal information. However, we wanted to see that the classification accuracy decreased as the number of samples gradually increased, but in our experiment, the classification performance increased as the number of samples increased.

In the tumer's classification experiment, we found that the same results were obtained when the real pathological image was used for training and the fake image generated by GAN was used for training. However, tumor classification is relatively easy to solve because it is more like a binary classification task. Therefore, further experiment should be conducted to see that the generated fake image is a perfect correspondence to the actual pathological image.

In both experiments, we have confirmed the possibility of what we intended, but it seems to be insufficient to fully confirm our intended results. In future work, we are going to experiment with more samples and images than ever before, and introduce more precise classification tasks to confirm our intended results.

CHAPTER IV

Conclusion

Through this thesis, we have researched about uses multi-scale information of images through pathological image analysis like image segmentation and image generation tasks.

In the the skin lesion segmentation research, we combined the results of different deeplearning networks that learned images on different scales. By using this again to train deeplearning models, the new network was designed with the goal of improving performance througha combined multiscale model rather than using a single model alone. The image segmentation method using multi-scale information has been proven to improve the accuracy of meaningful prediction results through skin lesion segmentation data. More research and model expansion needed to further study the accuracy, and further verification is needed to ensure that this accurate segmentation improves the diagnosis of skin cancer.

In the pathology image generation research, We generate a pathology image using SinGAN, which has a pipelined structure that utilizes multi-scale information using a single image, and applied it to two tasks. First, through the patient classification task, we classified fake pathological images generated from GAN into models overfitted on real pathological images, which resulted in about 30% reduced classification performance and increased patient privacy of pathological images. Secondly, we examined the possibility of using images generated from GAN as dataset for model training through tumor classification task. However, we think that these two experiments have been carried out with fewer samples, so we need more experiments after we have more samples.

References

[1] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680. iv, 2, 11

[2] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241. iv, 4, 8, 9

[3] Geoffrey Hinton, Oriol Vinyals, and Jeff Dean, "Distilling the knowledge in a neural network," *arXiv preprint arXiv:1503.02531*, 2015. iv, 5, 6

[4] Ilija Radosavovic, Piotr Dollár, Ross Girshick, Georgia Gkioxari, and Kaiming He, "Data distillation: Towards omni-supervised learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4119–4128. iv, 5, 6

[5] Noel CF Codella, David Gutman, M Emre Celebi, Brian Helba, Michael A Marchetti, Stephen W Dusza, Aadi Kalloo, Konstantinos Liopyris, Nabin Mishra, Harald Kittler, et al., "Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic)," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE, 2018, pp. 168–172. iv, 7

[6] Philipp Tschandl, Cliff Rosendahl, and Harald Kittler, "The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Scientific data*, vol. 5, pp. 180161, 2018. iv, 7

[7] Alec Radford, Luke Metz, and Soumith Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv preprint arXiv:1511.06434*, 2015. iv, 11

[8] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232. iv, 12

[9] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al., "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4681–4690. iv, 12

[10] Tamar Rott Shaham, Tali Dekel, and Tomer Michaeli, "Singan: Learning a generative model from a single natural image," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 4570–4580. iv, 3, 12, 13, 16, 18

[11] Oscar Geessink, Péter Bándi, Geert Litjens, and Jeroen van der Laak, "Camelyon17: Grand challenge on cancer metastasis detection and classification in lymph nodes," 2017. iv, 16, 17

[12] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 4, pp. 834–848, 2017. 1

[13] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia, "Pyramid scene parsing network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2881–2890. 1

[14] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee, "Enhanced deep residual networks for single image super-resolution," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2017, pp. 136–144. 1

[15] Yun Liu, Krishna Gadepalli, Mohammad Norouzi, George E Dahl, Timo Kohlberger, Aleksey Boyko, Subhashini Venugopalan, Aleksei Timofeev, Philip Q Nelson, Greg S Corrado, et al., "Detecting cancer metastases on gigapixel pathology images," *arXiv preprint arXiv:1703.02442*, 2017. 2

[16] Dayong Wang, Aditya Khosla, Rishab Gargeya, Humayun Irshad, and Andrew H Beck, "Deep learning for identifying metastatic breast cancer," *arXiv preprint arXiv:1606.05718*, 2016. 2

[17] Byungjae Lee and Kyunghyun Paeng, "A robust and effective approach towards accurate metastasis detection and pn-stage classification in breast cancer," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2018, pp. 841–850. 2

[18] Sanghun Lee, Sangjun Oh, Kyuhyoung Choi, and Sun Woo Kim, "Automatic classification on patient-level breast cancer metastases," . 2

[19] Heung-Il Suk, Seong-Whan Lee, Dinggang Shen, Alzheimer's Disease Neuroimaging Initiative, et al., "Hierarchical feature representation and multimodal fusion with deep learning for ad/mci diagnosis," *NeuroImage*, vol. 101, pp. 569–582, 2014. 2

[20] Rasool Fakoor, Faisal Ladhak, Azade Nazi, and Manfred Huber, "Using deep learning to enhance cancer diagnosis and classification," in *Proceedings of the international conference on machine learning*. ACM New York, USA, 2013, vol. 28. 2

[21] Jonathan Long, Evan Shelhamer, and Trevor Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440. 4

[22] Diederik P Kingma and Jimmy Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014. 7, 17

[23] François Chollet et al., "Keras," 2015. 7, 17

[24] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alexander A Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Thirty-First AAAI Conference on Artificial Intelligence*, 2017. 10

[25] Sergey Ioffe and Christian Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015. 12

[26] Martin Arjovsky, Soumith Chintala, and Léon Bottou, "Wasserstein gan," *arXiv preprint arXiv:1701.07875*, 2017. 12

[27] Xudong Mao, Qing Li, Haoran Xie, Raymond YK Lau, Zhen Wang, and Stephen Paul Smolley, "Least squares generative adversarial networks," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 2794–2802. 12

[28] Santiago Pascual, Antonio Bonafonte, and Joan Serra, "Segan: Speech enhancement generative adversarial network," *arXiv preprint arXiv:1703.09452*, 2017. 12

[29] Mehdi Mirza and Simon Osindero, "Conditional generative adversarial nets," *arXiv preprint arXiv:1411.1784*, 2014. 12

[30] Assaf Shocher, Shai Bagon, Phillip Isola, and Michal Irani, "Ingan: Capturing and remapping the "dna" of a natural image," in *International Conference on Computer Vision (ICCV)*, 2019, vol. 1, p. 2. 12

[31] Urs Bergmann, Nikolay Jetchev, and Roland Vollgraf, "Learning texture manifolds with the periodic spatial gan," in *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. JMLR. org, 2017, pp. 469–477. 12

[32] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron C Courville, "Improved training of wasserstein gans," in *Advances in neural information processing systems*, 2017, pp. 5767–5777. 14

[33] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778. 15, 16

[34] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer, "Automatic differentiation in pytorch," in *NIPS-W*, 2017. 17

Acknowledgement

I cannot believed that it has been already 2 years since I study for master degree. Without so many people around me, this thesis may not have been completed. First of all, I would like to thank to supervisor Professor Se Young Chun for his teaching about my study and research during master's course. Without his support, I would not have been able to learn about my major during my master's course. Also to my defence committee members: Professor Won-Ki Jung, Jeong-Hwan Jeon, I thank them for their advice. I would also like to acknowledge for my Lab mates: Hanvit Kim, DongWon Park, Jisoo Kim, KwanYoung Kim, Dongun Kang, Yonghyuk Seo, Termirlan Amangeldinov. And I would also like to thank my hometown friends including YoungHo Jeong, my friends at Chungwoon dormitory including NamKyu Park, my dormitory friends at Chungbuk including MyungHoon Jeong, and my 29TW friends including Jaeyoung Lee. Lastly, I sincerely thank to my family. In particular, I would like to thank JiSook Hong, who gave me a lot of advice and support during master's degree. And I thank my brother Seungjae Hong for going to graduate school soon.